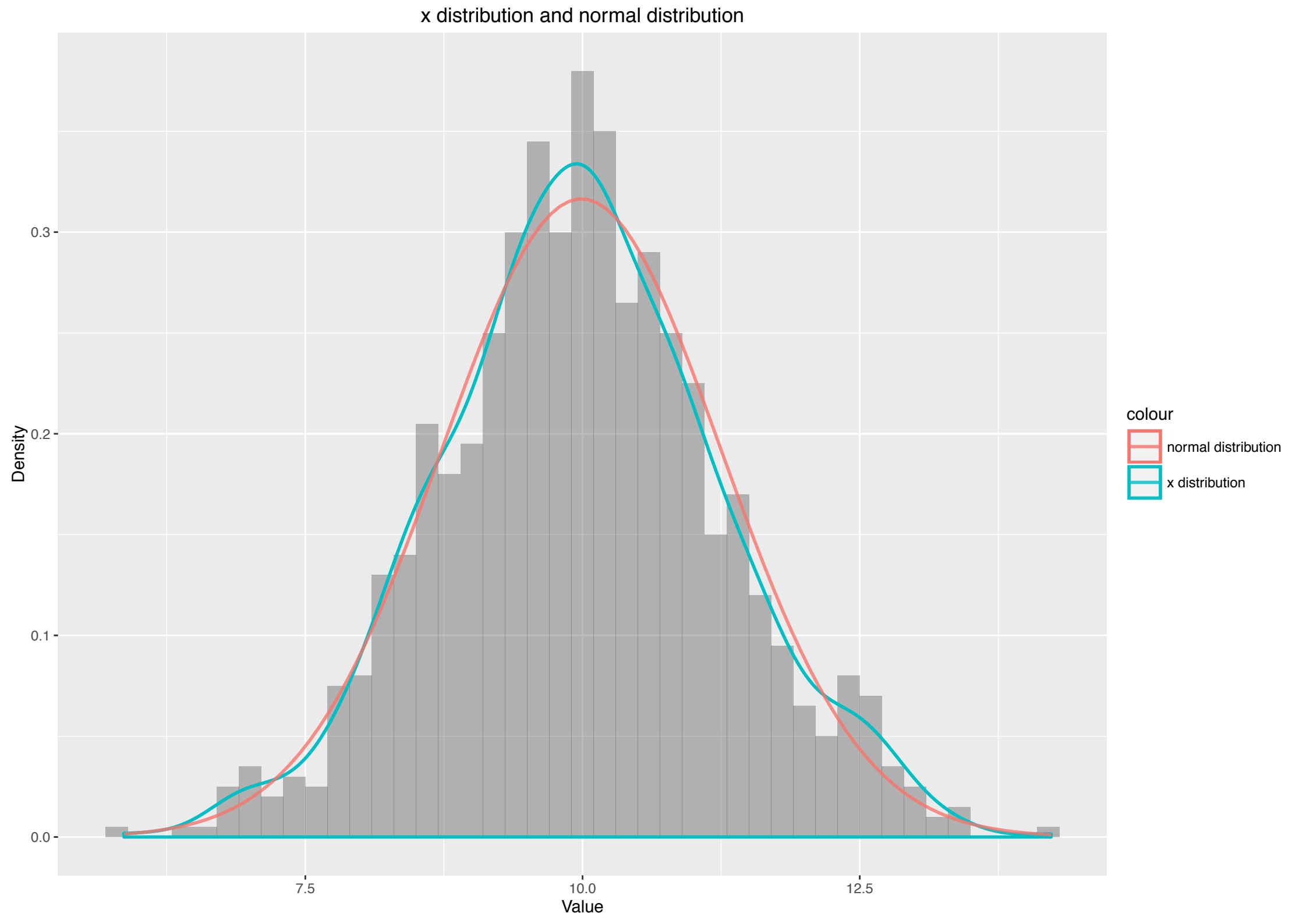
```
> ################## Gelman & Hill 2.2 ##################
> # data imput
> proportions_1908 <- c(.4777, .4875, .4859, .4754, .4874, .
4864, .4813, .4787, .4895, .4797, .4876, .4859)
> proportions_1909 <- c(.4857, .4907, .5010, .4903, .4860, .
4911, .4871, .4725, .4822, .4870, .4823, .4973)
> proportions_all <- c(.4777, .4875, .4859, .4754, .4874, .4864, .
4813, .4787, .4895, .4797, .4876, .4859,
+                      .4857, .4907, .5010, .4903, .4860, .4911, .
4871, .4725, .4822, .4870, .4823, .4973)
>
> # Compute the standard deviation of these proportions
> std_all <- sd(proportions_all)
> # observed sd
> std_all
[1] 0.006409724
>
> # compare to the standard deviation that would be expected
> # if the sexes of babies were independently decided with
> # a constant probability over the 24-month period
> n <- 3900
> p_exp <- mean(proportions_all)
> se_exp <- sqrt(p_exp*(1-p_exp)/n)
> # expected sd
> se_exp
[1] 0.008003121
> # se_exp > std_all
>
> # Is this difference statistically significant?
> df <- 23
> alpha <- 0.05
> upper <- qchisq(alpha/2., df)
> sd_upper <- sqrt(df*(std_all^2)/upper)
> lower <- qchisq(1 - (alpha/2.), df)
> sd_lower <- sqrt(df*(std_all^2)/lower)
> sd_lower
[1] 0.004981725
> sd_upper
[1] 0.008991309
> # As the observed standard deviation of 0.0064 is in the range
[0.005, 0.009]
> # we conclude that the difference is not signficant at 5%
```

```
> ################### Gelman & Hill 2.3 ##################
> library(ggplot2)
> n <- 20
> times <- 1000
> data <- runif(n*times, 0, 1)
> data <- matrix(data, nrow=1000, ncol=20, byrow = F)
> data <- data.frame(rowSums(data))
> std_data <- sd(as.matrix(data))
> mean_data <- mean(as.matrix(data))
>
> ggplot(data, aes(x=data)) +
+   geom_histogram(aes(y=..density..), binwidth = 0.2, alpha =0.4)
+
+   geom_density(lwd = 1, aes(col = 'x distribution'), alpha =
0.2) +
+   stat_function(fun=dnorm, args=list(mean=mean_data,
sd=std_data),
+                 lwd = 1, aes(col = 'normal distribution'), alpha
= 0.8) +
+   ggtitle("x distribution and normal distribution") +
xlab("Value") + ylab("Density")

> # Overall, the histogram of x and the normal distribution
overlap with each other. However, compared to the normal
distribution, the histogram of x shows a slight positive skew
trend and there is a small bumpon on the bottom right of the
histogram of x.
```


x distribution and normal distribution
Density
0.3
0.2
0.1
0.0
7.5
10.0
12.5
Value
colour
normal distribution
x distribution